

Real data analysis and power analysis using the developed package RIQITE

Devin Caughey, Allan Dafoe, Xinran Li, and Luke Miratrix

1 Install the RIQITE package

We first install the package from Github. The github page of the package (<https://github.com/li-xinran/RIQITE>) contains main functions with detailed explanation in R documentation, as well as a simple illustrating example. Below we will also use the package to analyze the real data sets in the paper.

```
devtools::install_github("li-xinran/RIQITE")
```

2 Testing monotonicity of an strumental variable

We get the data from Black et al. (2011), in which Table 1 cross-tabulates month of birth and school entry age (early/on-time/late) in terms of proportions and Table 3 reports the total number of subjects born in January or December (the “discontinuity subsample”): 104,023. We round the total number of subjects to 104,000, and, similar to Fiorini and Stevens (2014), we assume there are equal numbers of subjects born in January and December.

```
rm(list=ls())
library(dplyr)
library(ggplot2)
library(RIQITE)
set.seed(1)
x <- matrix(0, 2, 3, dimnames=list(Month=c("Dec", "Jan"),
                                     Entry=c("Early", "On-Time", "Late")))

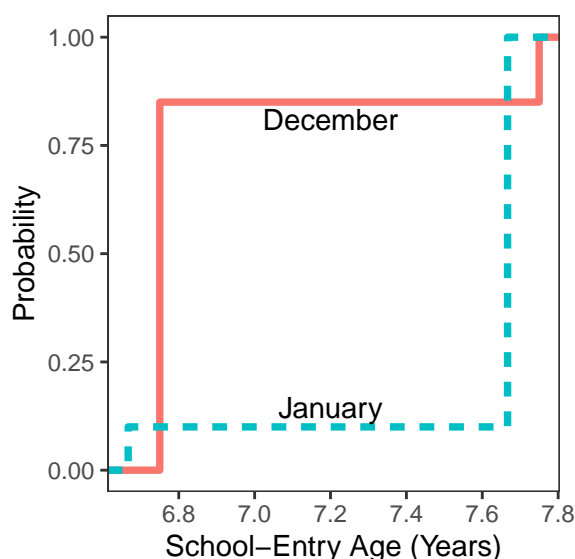
n <- 104000
x["Dec", "On-Time"] <- .85
x["Dec", "Late"] <- .15
x["Jan", "Early"] <- .10
x["Jan", "On-Time"] <- .90
df <- data.frame(BirthMonth = rep(c("Dec", "Jan"), each=n/2),
                 EntryTiming = c(rep("On-Time", n/2 * .85), rep("Late", n/2 * .15),
                                rep("Early", n/2 * .10), rep("On-Time", n/2 * .90)))
df <- df %>% mutate(EntryAge = NA,
                  EntryAge = ifelse(BirthMonth == "Dec" & EntryTiming == "On-Time", 6.75, EntryAge),
                  EntryAge = ifelse(BirthMonth == "Dec" & EntryTiming == "Late", 7.75, EntryAge),
                  EntryAge = ifelse(BirthMonth == "Jan" & EntryTiming == "On-Time", 23/3, EntryAge),
                  EntryAge = ifelse(BirthMonth == "Jan" & EntryTiming == "Early", 20/3, EntryAge))
data <- df[sample(c(1:n)), ] # randomly permute the ordering of units
table(data$BirthMonth, data$EntryTiming)
```

```
##
##      Early  Late On-Time
## Dec      0  7800  44200
```

```
## Jan 5200 0 46800
```

The following figure shows the empirical distribution functions of the school-entry age of December-born and January-born subjects. From the figure, most children started school at an older age if they were born in January rather than December.

```
ggplot(data, aes(EntryAge, color = BirthMonth, linetype=BirthMonth)) +
  stat_ecdf(geom="step", show.legend=FALSE, size=1.3) +
  annotate("text", x=c(7.2, 7.2), y=c(.145, .81), label=c("January", "December")) +
  scale_x_continuous(breaks=seq(6, 8, .2)) + theme_bw() +
  theme(plot.background = element_blank(), panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) + theme(plot.title=element_text(hjust=.5)) +
  labs(x="School-Entry Age (Years)", y="Probability")
```



We first perform the standard regression analysis assessing the relation between the instrument (birthday) and the school-entry age. The first-stage relationship is incredibly strong, with an average effect of -0.667 years and an F statistic of 106,256. This would conventionally be considered persuasive evidence of a valid instrument.

```
lm.fit <- lm(EntryAge ~ (BirthMonth=="Dec"), data)
print(paste(
  "esimated average effect is", round( as.numeric(lm.fit$coefficients[2]), digits = 3),
  "and F statistic is", round( as.numeric( summary(lm.fit)$fstatistic[1] ), digits = 0)
))
```

```
## [1] "esimated average effect is -0.667 and F statistic is 106256"
```

We then perform the randomization test for the bounded null $H_{\leq 0}$ on the monotonicity of the instrument. Specifically, we consider three randomization tests using the difference-in-means statistic, Wilcoxon rank sum statistic and Stephenson rank sum statistic with the parameter $s = 10$, respectively. We also conduct the classical Student's t-test. These produce Table 1 in the paper.

```
data$Z = as.numeric( data$BirthMonth == "Dec" )
data$Y = data$EntryAge
n <- length(data$Z)
m <- sum(data$Z)
nperm <- 10^4 # number of permutations for Monte Carlo approximation
Z.perm <- assign_CRE(n, m, nperm)
# difference-in-means
```

```

pval.dim <- pval_bound(Z=data$Z, Y=data$Y, c=0, method.list=list(name = "DIM"),
                      alternative = "greater", Z.perm=Z.perm, impute="control")
ci.dim<-ci_bound(Z=data$Z, Y=data$Y, alternative="greater", method.list=list(name="DIM"),
                 Z.perm=Z.perm, impute="control", alpha=0.10)

# Wilcoxon rank sum
pval.wilc <- pval_quantile(Z=data$Z, Y=data$Y, k=n, c=0, alternative = "greater",
                          method.list=list(name="Stephenson", s=2), Z.perm=Z.perm )
ci.wilc <- ci_quantile(Z=data$Z, Y=data$Y, k.vec = n, alternative = "greater",
                      method.list=list(name="Stephenson", s=2), Z.perm=Z.perm, alpha=0.10)

# Stephenson rank sum
pval.steph10 <- pval_quantile(Z=data$Z, Y=data$Y, k=n, c=0, alternative = "greater",
                             method.list=list(name="Stephenson", s=10), Z.perm=Z.perm )
ci.steph10 <- ci_quantile(Z=data$Z, Y=data$Y, k.vec = n, alternative = "greater",
                          method.list=list(name="Stephenson", s=10), Z.perm=Z.perm, alpha=0.10)

# classical Student's t-test
ttest = t.test(data$Y[data$Z==1], data$Y[data$Z==0], alternative = "greater")
pval.ttest = ttest$p.value
ci.ttest = round( ttest$conf.int[1], digits = 3)

# result for Table 1 in the paper
result = data.frame(test=c("diff-in-means", "Wilcoxon", "Stephenson 10", "t test"),
                    pval=c(pval.dim, pval.wilc, pval.steph10, pval.ttest),
                    lower = c(ci.dim, ci.wilc$lower, ci.steph10$lower, ci.ttest))

result

```

```

##          test pval  lower
## 1 diff-in-means    1 -0.669
## 2      Wilcoxon    1 -0.916
## 3 Stephenson 10    0  0.084
## 4          t test    1 -0.670

```

Finally, we infer the effect range. The result shows that the range of the effect of birth month on school-entry age is at least 1 year, indicating significant individual effect heterogeneity.

```

tau.max.lower = ci_quantile(Z=data$Z, Y=data$Y, k.vec = n, alternative = "greater",
                             method.list=list(name="Stephenson", s=10), Z.perm=Z.perm, alpha=0.05)$lower
tau.min.upper = ci_quantile(Z=data$Z, Y=data$Y, k.vec = 1, alternative = "less",
                             method.list=list(name="Stephenson", s=10), Z.perm=Z.perm, alpha=0.05)$upper
print(paste( "lower confidence limit for effect range is",
             max(0, tau.max.lower - tau.min.upper) ))

```

```
## [1] "lower confidence limit for effect range is 1.001"
```

```
save.image("BirthSchoolAge20220921.RData") # save the analysis result
```

3 Evaluating the effectiveness of professional development

We get the data from Heller et al. (2010). We remove units with missing outcomes, as well as units in Site 1 where there are no control outcomes.

```

rm(list=ls())
set.seed(2)
library(tidyverse)
library(RIQITE)
## DATA
dat = read_csv( "cleanTeacher.csv", na = "." )

```

```

dat = dplyr::select( dat, T.ID, Site, Tx, Per.1, Per.2, gain )
dat = mutate( dat, TxAny = ifelse( Tx == "D", 0, 1 ) )
# remove units with missing outcomes and units in Site 1 that has no control outcomes
dat = filter( dat, !is.na( gain ) & Site != "S1" )
dat = sample_n( dat, nrow(dat) )
n = nrow(dat)
table(dat$TxAny)

```

```

##
##      0      1
## 69 164

```

```

print(paste("estimated average effect is",
            round(lm(gain~TxAny, dat)$coefficients[2], digits = 2)))

```

```

## [1] "estimated average effect is 19.77"

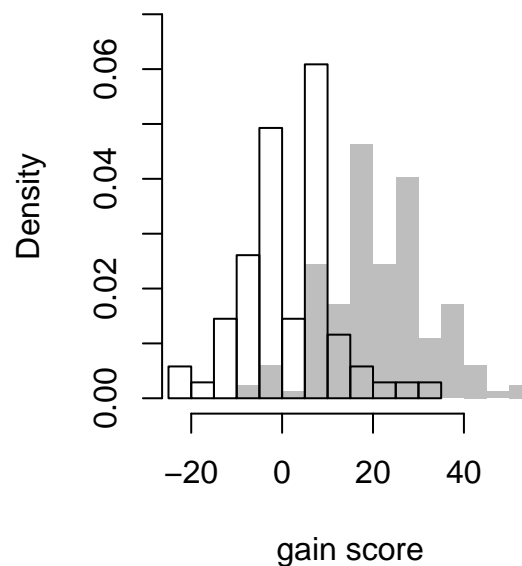
```

The following figure shows the histograms of the observed gain scores in the treatment and control groups.

```

hist(dat$gain[dat$TxAny==1], xlim = range(dat$gain), freq = FALSE, ylim = c(0, 0.07),
     col = "grey", border=FALSE, breaks = 20, main = NULL, xlab = "gain score")
hist(dat$gain[dat$TxAny==0], freq = FALSE, add = TRUE, breaks = 20, col = NULL)

```



We first test the bounded null hypothesis that all individual effects are non-positive.

```

nperm = 10^6
pval.steph10 <- pval_quantile(Z=dat$TxAny, Y=dat$gain, k=n, c=0, alternative = "greater",
                             method.list=list(name="Stephenson", s=10), nperm=nperm )
print(paste("The p-value is", pval.steph10))

```

```

## [1] "The p-value is 0"

```

We then conduct inference for all quantiles of individual treatment effects. The following figures show the 90% lower confidence limits for all quantiles of individual effects using Stephenson rank statistics with $s = 10$ and $s = 2$ (i.e., Wilcoxon rank), respectively.

```

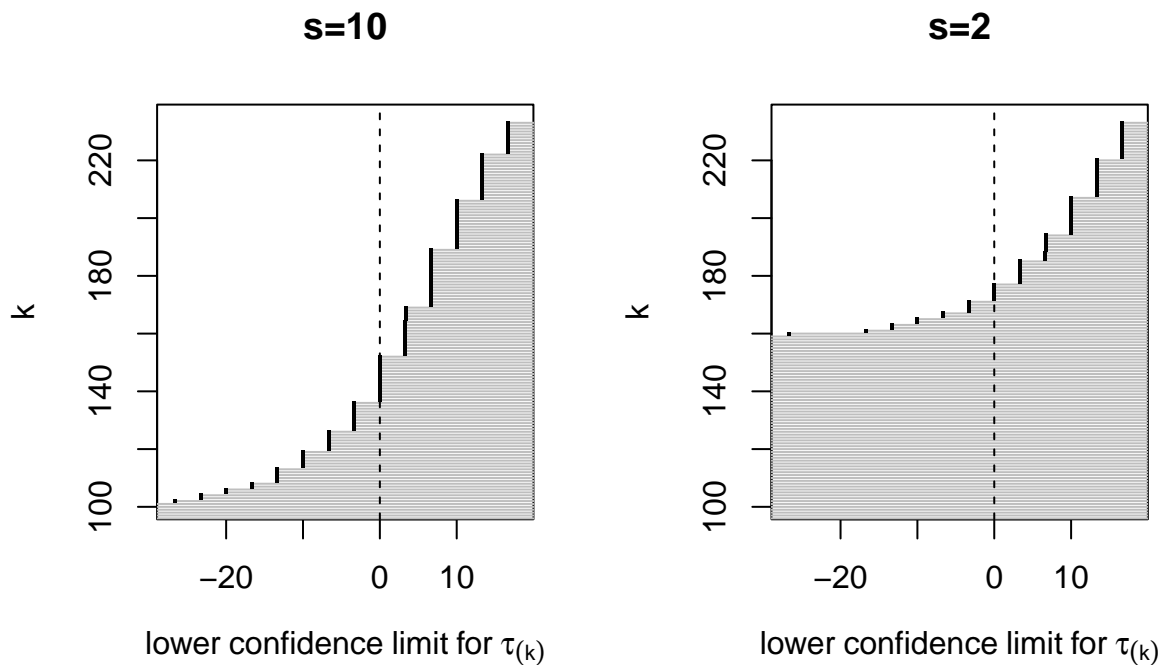
ci10 = ci_quantile( Z=dat$TxAny, Y=dat$gain, alternative="greater",
                   method.list=list( name="Stephenson", s=10 ), nperm=nperm, alpha=0.10 )
ci2 = ci_quantile( Z=dat$TxAny, Y=dat$gain, alternative="greater",

```

```

        method.list=list( name="Stephenson", s=2 ), nperm=nperm, alpha=0.10 )
par(mfrow = c(1, 2))
for(s in c(10, 2)){
  if(s==10){ci.limit = ci10$lower}
  if(s==2){ci.limit = ci2$lower}
  plot( NA, ylab="k", xlab=expression("lower"~"confidence"~"limit"~"for" ~tau[(k)]),
        ylim=c(102, 233)+c(-1, 1), xlim=range( ci.limit[ci.limit>=-Inf] )+c(-0.5, 1.5),
        main=paste0("s=", s) )
  for(k in 1:length(ci.limit)){
    lines( c( max( ci.limit[k], min(ci.limit[ci.limit>=-Inf]) - 100 ), max(ci.limit)+10),
          rep(k,2), col = "grey" )
  }
  points( ci.limit, c(1:length(ci.limit)), pch = ".", cex = 2 )
  abline(v = 0, lty = 2)
}

```



Finally, we infer the effect range.

```

tau.max.lower = ci_quantile( Z=dat$TxAny, Y=dat$gain, k.vec=n, alternative="greater",
                             method.list=list( name="Stephenson", s=10 ), nperm=nperm, alpha=0.05 )$lower
tau.min.upper = ci_quantile( Z=dat$TxAny, Y=dat$gain, k.vec=1, alternative="less",
                             method.list=list( name="Stephenson", s=10 ), nperm=nperm, alpha=0.05 )$upper
print(paste( "lower confidence limit for effect range is",
             max(0, tau.max.lower - tau.min.upper) ))

```

```
## [1] "lower confidence limit for effect range is 0"
```

```
save.image("Develop20220921.RData") # save the analysis result
```

4 Evaluating the effects of six-month nutrition therapy

We get the data for the homefood study from <https://doi.org/10.7910/DVN/38X3LX>. We download the SPSS Binary (Original File Format) there. The details of the study can be found at <https://clinicaltrials.gov>

/ct2/show/NCT03995303.

```
rm(list=ls())
set.seed(3)
library(foreign)
library(RIQUITE)
# source("power_simu_20220919.R")
# source("RI_Quantile_20220919.R")
# source("RI_Max_20220919.R")
dat = read.spss("Data file, 24.09.2021.sav", to.data.frame=TRUE)
```

We focus on the effect of the treatment on the increase of lean body mass (kg) measured before and after the trial. We exclude two units with missing outcomes, resulting in 52 treated units and 52 control units.

```
dat$change = dat$lean_mass_kg_1 - dat$lean_mass_kg_0
paste("There are", sum(is.na(dat$change)), "units with missing outcomes.")
```

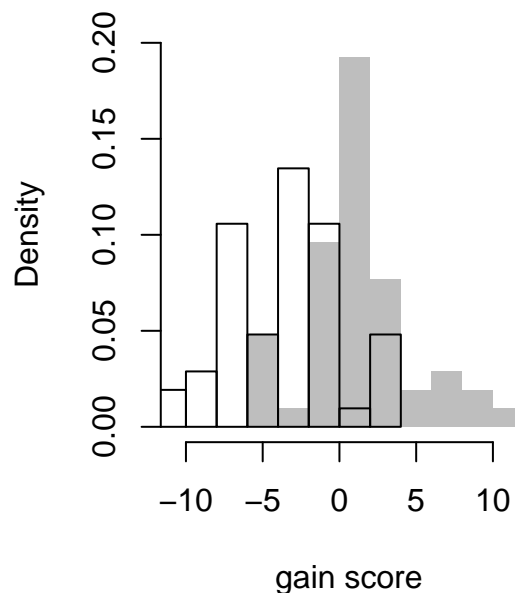
```
## [1] "There are 2 units with missing outcomes."
```

```
dat = dat[ !is.na(dat$change), ] # delete units with missing outcomes
n = nrow(dat)
dat = dat[sample(n), ] # randomly permute the order of units
table(dat$group)
```

```
##
##      control intervention
##          52          52
```

The following figure shows the histograms of the lean body mass changes in treated and control groups, respectively.

```
hist(dat$change[dat$group=="intervention"], xlim=range(dat$change), freq=FALSE,
      ylim=c(0,0.20), col="grey", border=FALSE, breaks = 10, main=NULL, xlab="gain score")
hist(dat$change[dat$group=="control"], freq=FALSE, add=TRUE, breaks=10, col=NULL)
```



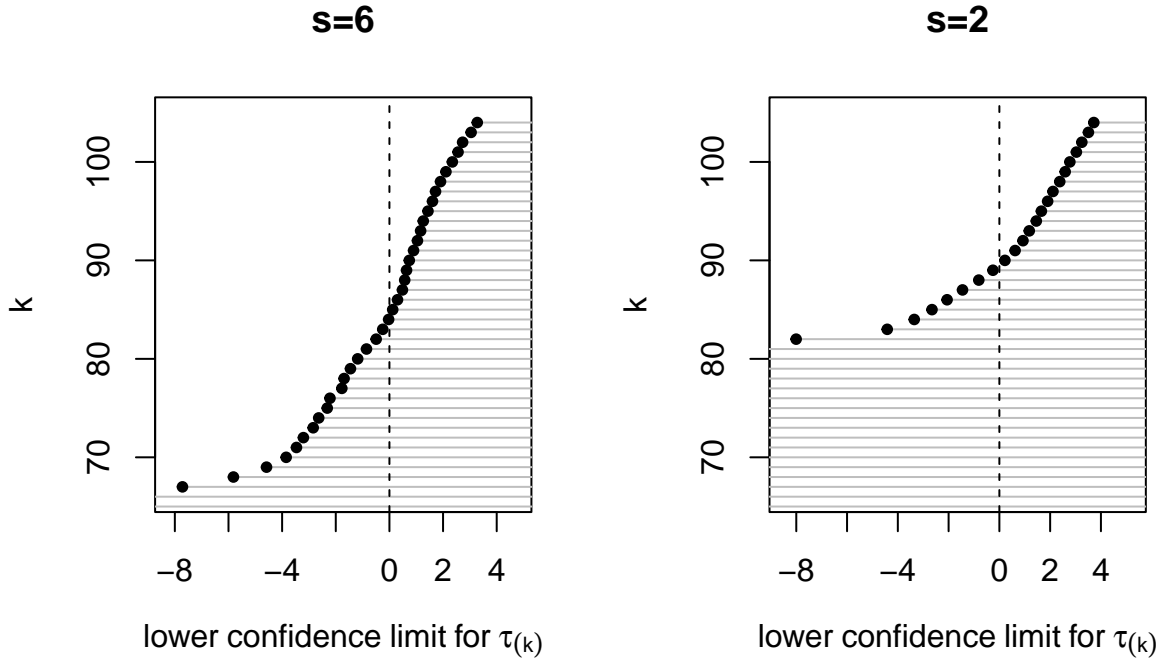
We infer the lower confidence limits for all quantiles of individual treatment effects. The following two figures show the 90% lower confidence limits for all quantiles of individual effects using Stephenson rank statistics with $s = 6$ and $s = 2$ (i.e., Wilcoxon rank), respectively

```

nperm = 10^6
ci6 = ci_quantile( Z=as.numeric(dat$group == "intervention"), Y=dat$change,
                  alternative="greater", method.list=list(name="Stephenson", s=6),
                  nperm=nperm, alpha=0.10 )
ci2 = ci_quantile( Z=as.numeric(dat$group == "intervention"), Y=dat$change,
                  alternative="greater", method.list=list(name="Stephenson", s=2),
                  nperm=nperm, alpha=0.10 )

par(mfrow = c(1, 2))
for(s in c(6, 2)){
  if(s==6){ci.limit = ci6$lower}
  if(s==2){ci.limit = ci2$lower}
  plot(NA, ylab="k", xlab=expression( "lower"~"confidence"~"limit"~"for" ~tau[(k)] ),
        ylim=c(n-37, n)+c(-1, 1)*1, xlim=range(ci.limit[ci.limit>-Inf])+c(-0.5, 1.5),
        main = paste0("s=", s))
  for(k in 1:length(ci.limit)){
    lines( c( max( ci.limit[k], min(ci.limit[ci.limit>-Inf]) - 100 ), max(ci.limit)+10),
           rep(k,2), col = "grey" )
  }
  points( ci.limit, c(1:length(ci.limit)), pch = 20 )
  abline(v = 0, lty = 2)
}

```



```

save.image("Homefood20220921.RData") # save the analysis result

```

5 Power comparison among different Stephenson rank sum statistics

We first give the data generating process. Here we consider the case where the control potential outcome follows the standard Gaussian distribution and the treatment potential outcome follows the Gaussian distribution with mean 1 and variance 2. This include the case where the individual effect follows Gaussian distribution mean 1 and variance 1, independent of the control potential outcome.

```
rm(list=ls())
library(RIQUITE)
gen_P0 <- function(n){
  mu0 = 0
  sd0 = 1
  mu1 = 1
  sd1 = 2
  Y1 = rnorm(n, mean = mu1, sd = sd1)
  Y0 = rnorm(n, mean = mu0, sd = sd0)
  return(data.frame(Y1 = Y1, Y0 = Y0))
}
```

We then conduct simulation over 100 simulated data sets, using Stephenson rank sum statistics with parameter s equal to 2, 6 and 10.

```
n = 120
iter.max = 500
nperm = 10^5
simu.result <- simu_power( gen=gen_P0, n=n, treat.prop=0.5, iter.max=iter.max,
  Steph.s.vec=c(2, 6, 10), alternative="greater", alpha=0.1,
  nperm=nperm, print.iter=100 )
```

```
## [1] "Monte Carlo approximation for null distributions completed..."
## [1] "Simulation for 100th iteration completed..."
## [1] "Simulation for 200th iteration completed..."
## [1] "Simulation for 300th iteration completed..."
## [1] "Simulation for 400th iteration completed..."
## [1] "Simulation for 500th iteration completed..."
```

```
save.image("PowerSimu20220921.RData") # save the simulation result
```

The code below compares the average lower confidence limits of $n(0)$, the number of units with effects greater than 0.

```
summary_power( result=simu.result, measure="CI_n(c)", alternative="greater", cutoff=0 )

##      s ci.mean
## 1  2    6.600
## 2  6   14.754
## 3 10   15.966
```

The code below compares the average lower confidence limits of $\tau_{(k)}$, the individual effect at rank k .

```
summary_power( result=simu.result, measure="CI_tau(k)", alternative="greater",
  cutoff=NULL, k=n )
```

```
##      s ci.mean
## 1  2    0.5990
## 2  6    1.3580
## 3 10    1.5995
```

The code below compares the power for testing the null hypothesis $H_0 : \tau_{(k)} \leq 1$, with k equals the sample size n .

```
summary_power( result=simu.result, measure="test_tau(k)", alternative="greater",
  cutoff=1, k=n )
```

```
##      s power.mean
## 1  2         0.102
```


##	2	6	0.836
##	3	10	0.942

References

- Black, S. E., Devereaux, P. J., and Salvanes, K. G. (2011). Too young to leave the nest? the effects of school starting age. *Review of Economics and Statistics*, 93:455–467.
- Fiorini, M. and Stevens, K. (2014). Assessing the Monotonicity Assumption in IV and Fuzzy RD designs. Working papers, University of Sydney, School of Economics.
- Heller, J. L., Shinohara, M., Miratrix, L., Hesketh, S. R., and Daehler, K. R. (2010). Learning science for teaching: Effects of professional development on elementary teachers, classrooms, and students. *Proceedings from Society for Research on Educational Effectiveness*.